

TALLINNA ÜLIKOOL

Haapsalu kolledž

Rakendusinformaatika õppekava

Hanna Kalavus

HAAPSALU KOLLEDŽI KIRJALIKE TÖÖDE VORMINDUSE  
AUTOMAATKONTROLI VÕIMALIKKUS

Diplomitöö

Juhendaja: Priidu Paomets MSc

Haapsalu 2023

## SISUKORD

MÕISTED JA LÜHENDID .....	4
SISSEJUHATUS.....	5
1. METOODIKA .....	7
2. UURINGU TULEMUSED.....	9
2.1. Sarnased lahendused .....	9
2.2. Üliõpilaste küsitluse tulemused.....	10
2.2.1. Vastanute kursus ja eriala.....	10
2.2.2. Haapsalu kolledži õpingute ajal esitatud kirjalike tööde arv.....	10
2.2.3. Haapsalu kolledži õpingute ajal esitatud kirjalike tööde tüübid .....	11
2.2.4 Vormindusnõuetele vastavate kirjalike tööde esitamise sagedus.....	12
2.2.5. Peamine tekstitöötlusprogramm ja dokumendi esitamise formaat.....	13
2.3. Õppejõudude küsitluse tulemused.....	13
2.3.1. Õppejõudude erialad .....	13
2.3.2. Kirjalike tööde vorminduse olulisus, mugavus ja ajakulu .....	13
2.3.3. Kirjalike tööde lugemise/kontrollimise keerukus .....	14
2.3.4. Võimaliku automaatkontrolli rakenduse plussid.....	14
2.3.5. Võimaliku automaatkontrolli rakenduse miinused .....	14
2.3.6. Ootused võimalikule automaatkontrolli rakendusele .....	15
3. FAILIFORMAADID.....	16
3.1. PDF.....	16
3.2. DOCX.....	17
3.3. Järeldus .....	17
4. VORMINDUSNÕUDED .....	19
4.1. Nõuete kõrvutamine dokumendi parameetritega.....	20
4.1.1. Faili lahtipakkimine.....	20
4.1.2. Faili lugemine.....	20
4.1.3. Faili kõrvutamine .....	21
4.2. Olulised tähelepanekud .....	21
4.2.1. Mõõtühikud .....	21
4.2.2. Stiilihierarhia .....	21

5. PARAMEETRID.....	24
6. VÕIMALIKUD LAHENDUSED KONTROLLI AUTOMATISEERIMISEKS .....	26
6.1. Töövoog .....	26
6.2. Võimalik realiseerimine .....	27
7. TULEMUSTE ANALÜÜS JA JÄRELDUSED.....	28
KOKKUVÕTE.....	29
SUMMARY .....	30
ALLIKAD .....	31
LISA 1. ÜLIÕPILASTE KÜSIMUSTIK	
LISA 2. ÕPPEJÕUDUDE KÜSIMUSTIK	
LISA 3. LEHEKÜLJE KUJUNDUS	
LISA 4. PEALKIRJAD	
LISA 5. TABELID, JOONISED, ILLUSTRATSIOONID JA FOTOD	

## MÕISTED JA LÜHENDID

**DOCX** – Microsofti patenteeritud dokumendivorming (Microsoft, 2023).

**DXA** - Microsoft Wordi standardne mõõtühik, vastab ühele kahekümnendikule punktist ( $1 \text{ dxa} = \frac{1}{20} \text{ pt}$ ) (Microsoft, 2023).

**JPG** - digitaalpiltide pakkimise standard, mis on mõeldud vaid digitaalfotograafia jaoks (Davis, 2023).

**Office Open XML** - Microsofti poolt välja töötatud kokkupakitud XML-põhine failiformaat (Dick, s.a.).

**PDF** (*Portable Document Format*) - Adobe poolt loodud kaasaskantav dokumendivorming. PDF sisaldab tekste, pilte ja muid elemente ja seda kasutatakse vormindatud tekstisisu kuvamiseks (ISO, 2020).

**Plugin** (*pistikprogramm*) - tarkvaramoodul, mis lisab mõnele teisele tarkvarasüsteemile võimalusi (George, 2021).

**PNG** (*Portable Network Graphics*) - rastergraafika failivorming, mis on internetis enim levinud (Adobe, s.a.).

**XML** - tekstipõhine vorming struktureeritud teabe esitamiseks (Library Of Congress, 2019).

**ZIP** – failivorming, mida kasutatakse ühe või mitme faili kokkusurumiseks ühte asukohta (Yakovenko, 2016).

## SISSEJUHATUS

Käesoleva diplomitöö teemaks on Haapsalu kolledži kirjalike tööde vorminduse automaatkontrolli võimalikkus. Teema on aktuaalne, sest kirjalikke töid esitatakse alates esimesest semestrist, kuid täna on võimalik kirjalike tööde vorminduse kontrolli teostada vaid manuaalselt ning see on ajakulukas. Tööde vormindamine võib osutuda tülikas, mistõttu ei pöörata sellele suuremat tähelepanu. Vormindus on aga suur osa kirjalikust tööst. Kontrolli teostav rakendus võiks olla abiks harjumuse tekkimisel ja pikemas perspektiivis on tudengitel lihtsam kirjalikke töid vormindada ning õppejõududel neid lugeda.

Hüpoteesiks on, et arendatav rakendus hoiaks märkimisväärselt kokku õppejõudude aega, mis kulub kirjalike tööde vorminduste kontrollimisele. Tänu sellele saaksid ka tudengid kiiremini oma töödele tagasisidet. See tähendab, et kogu kirjalike tööde protsess võiks kiirenedada.

Diplomitöös proovitakse leida läbi uurimisküsimuste ning -tegevuste kinnitust esitatud hüpoteesile. Lisaks soovitakse tööga leida võimalikku lahendust probleemile läbi teoreetiliste aluste.

Diplomitöö käsitleb järgmist probleemküsimust:

Kas ja kuidas on võimalik vähendada õppejõududel kirjalike tööde vorminduse kontrollile kuluvat aega?

Töö eesmärgiks on uurida kirjalike tööde vorminduse automaatkontrolli võimalikkust. Pakutud lahenduseks nimetatud probleemile on rakendus, mis kontrollib kirjalike tööde vormindust vastavalt Haapsalu kolledži kirjalike tööde vorminduse juhendile. Lahendus võiks vähendada õppejõududel kirjalike tööde vorminduse kontrollile kuluvat aega, sest kaotaks protsessist manuaalse faktori.

Uurimisprobleemist ja eesmärgist lähtuvalt esitati järgmised uurimisküsimused:

- Millised sarnased lahendused on täna juba olemas ja kas neid saaks rakenduse jaoks kasutada?
- Kuidas saab kirjalikku tööd vastu nõudeid kontrollida ja milliseid parameetreid peaks rakendus jälgima?

- Milline on peamine failiformaat, milles tudengid oma töid esitavad ning milliseid failiformaate on võimalik kontrollida?
- Kas ja kuidas saaksid pakutavast lahendusest kasu nii õppejõud kui ka tudengid?

Uurimisküsimustele vastuste leidmiseks püstitati järgmised ülesanded:

- Uurida, kas ja millised sarnased lahendused on täna olemas ning kas neid saaks kasutada;
- Koostada küsimustikud saamaks ülevaadet kirjalike tööde tänasest olukorrast ja peamistest kasutusel olevatest formaatidest ja rakendustest;
- Uurida, milliseid failiformaate on võimalik kontrollida ja milliseid on otstarbekas kontrollida;
- Harutada lahti valitud failiformaadid ja kõrvutada parameetreid Haapsalu kolledži kirjalike tööde juhendiga.

Diplomitöö koosneb seitsmest peatükist, sissejuhatusest ja kokkuvõttest. Peatükkidest esimene kirjeldab uurimismeetodeid, teine uurimistulemusi. Kolmandas peatükis lahatakse failiformaate ning selle tulemusena on neljandas peatükis kirjeldatud valituks osutunud failiformaadi kõrvutamist kirjalike tööde nõuetega. Viiendas peatükis selgitatakse lahti töö põhiosana valminud parameetrite tabel. Kuuendas peatükis kirjeldatakse automaatkontrolli võimalikke etappe ning võrreldakse erinevaid potentsiaalseid lahendusi rakenduse realiseerimiseks. Seitsmes peatükk sisaldab tulemuste analüüsi ja järeldusi. Töö lisadest leiab küsitluste ankeedid ning parameetrite tabelid.

Kõik töös kasutatavad joonised, tabelid ja koodinäited on autori enda loodud kui pole viidatud teisiti.

## 1. METOODIKA

Järgnevas peatükis kirjeldatakse uurimuses kasutatavaid uurimis- ja andmekogumismeetodeid. Antakse ülevaade, mis tingimustel on koostatud valimid ja kuidas uurimus on läbi viidud. Uuringut planeerides lähtuti eesmärgist kaardistada kirjalike töid puudutav hetkeolukord ning koguda sisendeid võimaliku automaatkontrolli loomiseks.

Käesoleva diplomitöö esimene osa, üliõpilaste küsimustik, on kvantitatiivne. Selle jaoks on kasutatud kaardistusuuringut.

Kaardistusuuringuga (*survey*-tüüpi uuring) püütakse selgitada tegelikkuses valitsevaid tingimusi ja asjaolusid. Urija on sellisel juhul passiivne andmete koguja, kes rahuldub nähtuste loomuliku varieeruvuse registreerimisega, püüdmata olukordade või uuritavatega manipuleerida. (Õunapuu, 2014)

Uuringuna sai loodud üliõpilastele küsimustik. Kuna üliõpilastele koostatud küsimustiku peamine eesmärk oli saada konkreetset ülevaadet kasutatavate rakenduste ning enimlevinud failiformaatide kohta, siis oli küsimustik jäigalt struktureeritud valikvastustega.

Kuna nimetatud teema puudutab kogu Haapsalu kolledžit akadeemilist pere, siis sai valimiks kogu kolledži üliõpilaste koosseis. Küsitluse eesmärgiks oli jõuda võimalikult paljude tudengiteni, mistõttu valiti küsimustiku koostamiseks Google loodud küsitluste haldustarkvara *Google Forms* ja levitati seda üliõpilaste seas meili teel. Üliõpilasteni jõudmisel oli abiks Haapsalu kolledži õppenõustaja- ja spetsialist, kes nimetatud küsimustikku kooli meililistidele saatis.

Diplomitöö teine osa, õppejõudude küsimustik, on kvalitatiivne ning selle jaoks sai õppejõududele loodud avatud küsimustega küsitlus. Küsitluse põhieesmärgiks oli saada sisendeid võimaliku automaatkontrolli jaoks. Nimetatud küsitlus oli ülesehituselt sarnane avatud küsimustega intervjuuga, lisaks oli kasutatud ka valikvastuseid.

Õppejõudude küsitluse valimiks sai hulk Haapsalu kolledži erinevate kursuste õppejõude, kes populatsiooni arvamuse kataksid. Sisendi sobiva valimi jaoks tuli Haapsalu kolledži õppenõustaja- ja spetsialistilt. Vastanute arvust olulisem oli selle küsitluse puhul kõikide

Haapsalu kolledži erialade katvus, sest eesmärgiks oli mõista tervikpilti mitmest eri vaatenurgast. Ka õppejõudude küsitlus koostati *Google Forms*is.

Lisaks eelnimetatud meetoditele on antud töös analüüsitud ka kirjandust. Failiformaatide kohta informatsiooni kogumiseks on eritletud enimkasutatavaid failiformaate. Nõuete aluseks on võetud Haapsalu kolledži kirjalike tööde juhend ja nende kõrvutamiseks parameetritega on kasutatud ametlikke dokumentatsioone. Peamisteks kirjalikeks allikateks on elektroonilised artiklid, dokumentatsioonid ja standardid.



## 2. UURINGU TULEMUSED

Järgnev peatükk analüüsib diplomitöö käigus läbiviidud uurimuse tulemusi. Peatükis kirjeldatakse, kuidas sarnaseid lahendusi otsiti ja millistele järeldustele jõuti. Antakse ka ülevaade, milliseid küsitlusi failiformaadi valiku ning automaatkontrolli sisendi jaoks kasutati.

### 2.1. Sarnased lahendused

Analüüsiti sarnaste rakenduste võimalikku olemasolu. Esmalt alustati otsimist eesti keeles. Google otsingumootoris kasutati selliseid märksõnu nagu „uurimistöö vorminduse automaatne kontroll“, „faili vorminduse kontroll“, „failiformaadi kontroll“ jms. Kiirelt selgus aga, et eestikeelseid juhtnööre, mis edasi viiks, ei ole. Alustati ingliskeelsete päringute teostamisega.

Prooviti minna veidi konkreetsemaks ning valiti otsinguks üks kindlat failiformaati kasutav tekstitöötlusprogramm Word. Valik tehti esmalt autori isikliku kogemuse põhjal: senini on tööde vormindamiseks kasutatud nimelt just Wordi ning selle dokumendiformaati DOCX. Kasutati märksõnu „word document control“, mille otsingutulemus viitas vaid Wordi dokumendi sisukontrolli loomisele (nt. ühtne lubatud kuupäeva formaat). Liiguti edasi ja uuriti märksõnade „MS Word formatting“ järgi. Antud otsingu vastes seletati, mida üldse Wordi dokumendi vormindus tähendab ning mis sorti vormindust on võimalik kasutada. Katsetati veel mitmeid sarnaseid märksõnu, mis poleks liiga spetsiifilised, aga ühel või teisel viisil juhtnööridena edasi viiks. Soovitud vasteid ei leitud.

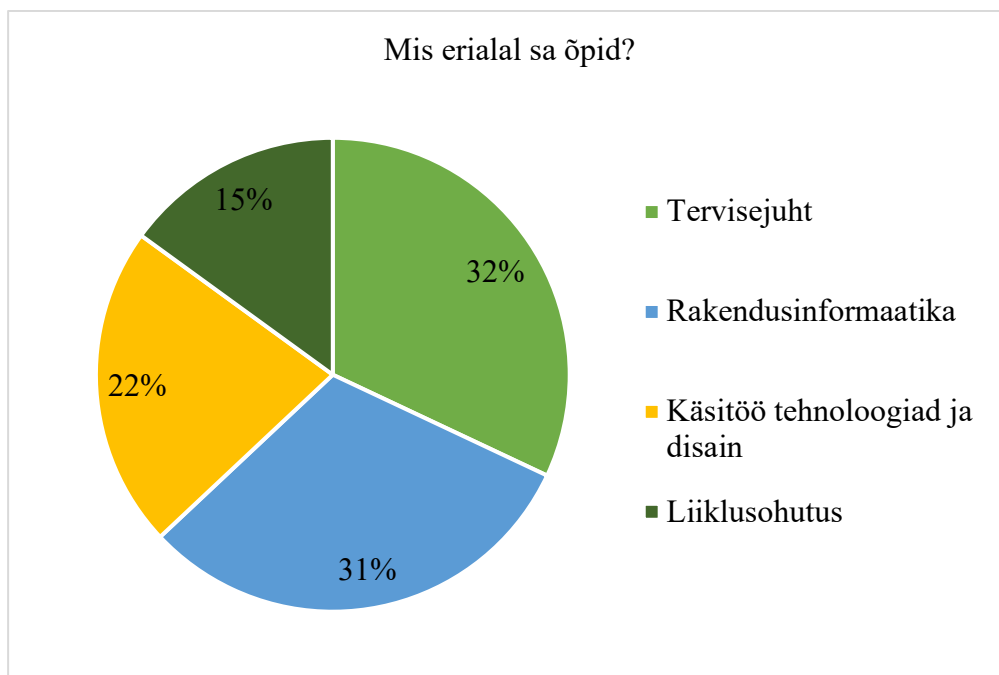
Olles veendunud, et taoline vorminduse automaatne kontroll ei ole laialtlevinud, tuli hakata spetsiifilisemalt lähenema. Järgmiseks otsustati uurida, millest üks Wordi dokument koosneb. Faililaiendi valik tehti nüüd juba ülikoolis kasutatava ja õpetatava tekstitöötlusprogrammi järgi. Märksõnad „DOCX formatting“ viisid juba konkreetset failiformaadi ülesehituse juurde, millest ka peatükis 3.1. juttu tuleb. Et aga teada saada, millise failiformaadi kohta süvitsi edasi peaks uurima, tuli esmalt välja selgitada, millised failiformaadid Haapsalu kolledžis kasutusel on. Selleks loodi küsitlused, mille tulemusi järgnevalt tutvustatakse.

## 2.2. Üliõpilaste küsitluse tulemused

Üliõpilaste küsimustik koosnes seitsmest küsimusest. Haapsalu kolledži üliõpilastelt uuriti nende senist kogemust kirjalike tööde kirjutamise, vormindamise ning esitamisega. Vastanute seas oli üliõpilasi kõikidelt Haapsalu kolledži erialadelt ja kursustelt. Küsimustikule vastas 100 üliõpilast.

### 2.2.1. Vastanute kursus ja eriala

Enim vastanuid (39%) oli kolmanda kursuse tudengite seas. Teiselt kursusel oli vastanuid 32% ning esimeselt 29%. Erialade raames oli enim vastanuid tervisejuhtide seas, siis rakendusinformaatika, käsitöetehnoloogiad ja disain ning liiklusohutus erialalt. Osakaal numbrites jagunes vastavalt 32%, 31%, 22%, 15% (joonis 1).



Joonis 1. Üliõpilaste jaotus erialade vahel

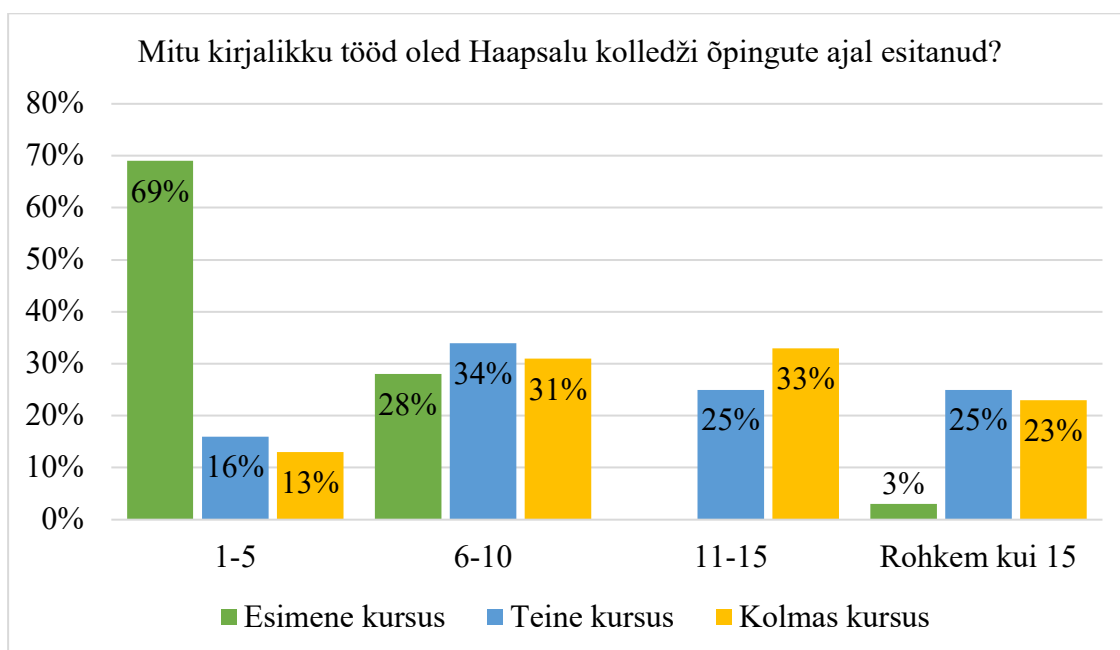
### 2.2.2. Haapsalu kolledži õpingute ajal esitatud kirjalike tööde arv

Saamaks ülevaadet kirjalike tööde mahust ja nende vormindamise aktuaalsusest uuriti üliõpilastelt, kui palju kirjalikke töid on nad oma Haapsalu kolledži õppetöö jooksul esitanud.

Küsitlus viidi läbi sügissemestri lõpus. See tähendab, et õpilased vastasid õppeaasta esimese semestri seisu silmas pidades.

Selgus, et 69% esimese kursuse üliõpilastest on esitanud esimese semestri lõpuks 1-5 kirjalikku tööd (joonis 2). 28% esimese kursuse vastanutest on esimese semestri lõpuks esitanud juba 6-10 kirjalikku tööd.

Kolmandaks õppeaastaks on aga 33% õpilastest esitanud 11-15 kirjalikku tööd, 23% rohkem kui 15 kirjalikku tööd ning 31% 6-10 kirjalikku tööd (joonis 2).

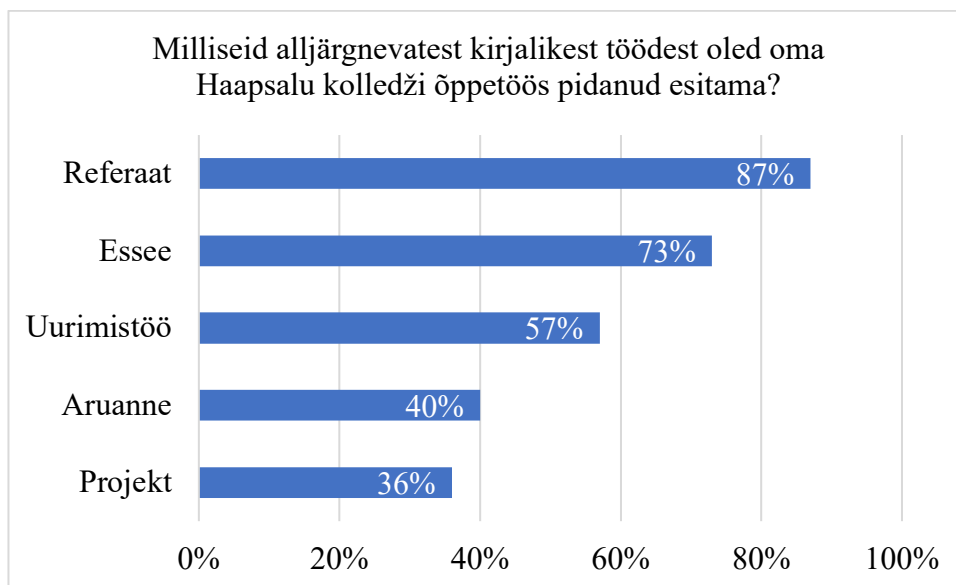


Joonis 2. Haapsalu kolledži õpingute ajal esitatud kirjalike tööde arv

### 2.2.3. Haapsalu kolledži õpingute ajal esitatud kirjalike tööde tüübid

Et saada teada kui paljud üliõpilaste kirjalikest töödest sarnanevad lõputöö struktuuriga ning vajavad täiemahulist vormindust, siis uuriti vastanutelt, milliseid kirjalikke töid on nad pidanud Haapsalu kolledži õppetöös esitama.

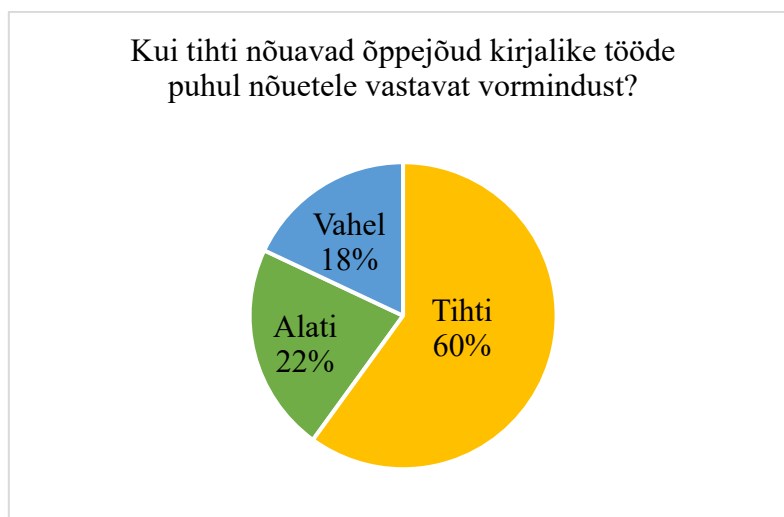
Selgus, et kõige rohkem on pidanud üliõpilased õppetöös esitama referaate – 87 vastanut (joonis 3). Viie populaarsema vastuse hulgas oli vaid üks kirjalik töö, mis struktuuri poolest lõputööst erineb: essee (joonis 3). Nimelt ei ole esseel reeglina tiitellehte, sisukorda, alapeatükke jne.



Joonis 3. Haapsalu kolledži õpingute ajal esitatud kirjalike tööde tüübid

#### 2.2.4 Vormindusnõuetele vastavate kirjalike tööde esitamise sagedus

Üliõpilastel paluti hinnata, kui tihti nende arvates õppejõud kirjalike tööde puhul nõuetele vastavat vormindust nõuavad (joonis 4). Vastused jagunesid kolmeks:



Joonis 4. Vormindusnõuetele vastavate kirjalike tööde esitamise sagedus

### **2.2.5. Peamine tekstitöötlusprogramm ja dokumendi esitamise formaat**

Vastanutelt uuriti ka, millist tekstitöötlusprogrammi nad peamiselt kirjalike tööde kirjutamisel ja vormindamisel kasutavad ja mis formaadis nad oma kirjalikke töid esitavad. Selgus, et 88% üliõpilastest kasutab kirjalike tööde kirjutamisel ja vormindamisel Microsoft Office Wordi, kuid vaid 13% esitab töid .docx formaadis. Enamik üliõpilastest, 87%, esitab kirjalikke töid .pdf formaadis.

## **2.3. Õppejõudude küsitluse tulemused**

Saamaks kirjalike töödega seonduvast sarnast sisendit ka õppejõudude vaatenurgast, uuriti Haapsalu kolledži erinevate erialade õppejõududelt, kas nende jaoks on kirjalike tööde puhul vormindus oluline, millised on kirjalike tööde lugemise/kontrollimise kitsaskohad ja mis on nende arvamused võimaliku automaatkontrolli kohta. Õppejõudude jaoks loodi küsitlus, mis koosnes kaheksast küsimusest.

### **2.3.1. Õppejõudude erialad**

Küsitlusele vastas 6 õppejõudu, kes katsid kõik erialad ära. Kaks õppejõudu annavad aineid tervisejuhtide erialale, kaks rakendusinformaatika erialale ning kaks üleülikooliliselt kõikidele erialadele.

### **2.3.2. Kirjalike tööde vorminduse olulisus, mugavus ja ajakulu**

Esmalt uuriti kui oluliseks õppejõud kirjalike tööde vormindust peavad. Kõik vastanud õppejõud peavad väga oluliseks, et neile esitatav kirjalik töö oleks nõuetele vastavalt vormindatud. Neli õppejõudu kuuest peab kirjalike tööde vorminduse kontrollimist (vigade kommenteerimine, õpilasele tagasisaatmine) ajakulukaks. Sama hulk vastanuid peab vormindusnõuetele mittevastava töö lugemist ebamugavaks.

### 2.3.3. Kirjalike tööde lugemise/kontrollimise keerukus

Õppejõudude käest küsiti, mis teeb kirjalike tööde lugemise/kontrollimise nende jaoks enim keeruliseks. Vastustest käis enim läbi tõsiasi, et kirjalike tööde lugemine ning kontrollimine on erinevatel põhjustel ajakulukas. Näiteks kui töö on lohakalt vormindatud, kulutab õppejõud tagasisidele rohkem aega. Ka suur esitatud tööde arv kasvatab ajamahtu.

### 2.3.4. Võimaliku automaatkontrolli rakenduse plussid

Õppejõududel paluti ette kujutada rakendust, mis aitaks üliõpilastel kirjaliku töö vormindusvigadele tähelepanu pöörata. Kirjeldusele tuginedes paluti õppejõududel rakenduse plusse/miinuseid analüüsida ja mõtteid jagada. Rakenduse kirjeldus õppejõududele oli järgnev: „Kujuta ette, et on olemas rakendus, mille abil üliõpilased enne kirjaliku töö esitamist selle vormindust saavad kontrollida. Rakendus annaks näiteks märku kui õpilase kontrollitava töö veerised ja teksti joondus ei vasta Haapsalu kolledži kirjalike tööde nõuetele. Pange tähele, et rakendus pööraks vaid tähelepanu vormindusvigadele, kuid ei parandaks neid.“

Esmalt küsiti, mis võiksid olla sellise rakenduse plussid. Vastustest on toodud välja tähtsamad märksõnad:

- Rakendus teeks kindlasti tudengitele töö kergemaks;
- Õppejõule annaks see kindlasti aja kokkuhoidu;
- Tööd oleks ühtses stiilis ning paremini loetavad;
- Tagasiside andmisel ei peaks vormindamist eraldi puudutama;
- Õppejõuni jõuaks korrektselt vormindatud tööd;
- Üliõpilased saaksid vormindusnõudes paremini selgeks.

### 2.3.5. Võimaliku automaatkontrolli rakenduse miinused

Küsiti ka, mis võiksid olla sellise rakenduse miinused ning esmajoones osati tähelepanu pöörata järgnevale:

- Programmil ei pruugi olla piisavalt andmeid või need ei ole piisavalt täpsed, et kontroll toimiks;

- Vormindamise reeglid võivad ajas muutuda;
- Üliõpilased võivad muutuda laisemaks.

### **2.3.6. Ootused võimalikule automaatkontrolli rakendusele**

Õppejõududel paluti lühikese rakenduse kirjelduse põhjal mõelda ka, mida nad selliselt rakenduselt ootaksid. Välja on toodud mõned mõtted:

- Ootan, et esitatud kodutööde vorminduse kvaliteet paraneks ja saaksin keskenduda sisulise tagasiside andmisele;
- Vea leidmisel võiks rakendus tudengile suuniseid anda, kuidas vormindust parandada.
- Kuna vormindusnõuded muutuvad, siis peaks rakendus olema lihtsasti kohandatav erinevatele nõuetele;
- Õpilasel võiks olla kindlam tunne oma töö ära saata, kui ta teab, et vähemalt vorminduse suhtes ei ole midagi kahe silma vahele jäänud.

### 3. FAILIFORMAADID

Järgnevas peatükis kirjeldatakse üliõpilaste poolt dokumentide redigeerimiseks ja edastamiseks enim kasutatud failiformaate. Võrreldakse nende ülesehitust ning langetatakse analüüsijärgselt parim võimalik valik automaatkontrolli teostamiseks.

Küsitlusest üliõpilastele selgus, et 88% tudengitest kasutab kirjalike tööde kirjutamisel ja vormindamisel Microsoft Office Wordi ning 87% üliõpilastest esitab oma töid .PDF formaadis. Sellest tulenevalt võetakse luubi alla kaks failiformaati: PDF ja DOCX.

#### 3.1. PDF

PDF (*Portable Document Format*) on Adobe poolt loodud kaasaskantav dokumendivorming. PDF sisaldab tekste, pilte ja muid elemente ja seda kasutatakse vormindatud tekstisisu kuvamiseks. PDF-failid konverteeritakse tavaliselt muudest tekstivormingutest, et muuta failid universaalselt ühilduvateks. (ISO, 2020)

Adobe eesmärk PDF'i loomisel oli võimaldada igaühel jäädvustada dokumente mistahes rakendusest, saata nende dokumentide elektroonilisi versioone kõikjale ja vaadata ning printida neid mistahes arvutist. (ISO, 2020)

PDF-fail koosneb vektor- ja rastergraafikast ning tekstide kompositsioonist stabiilsel kujul, mis tähendab, et seda saab avada mistahes tugitarkvaraga ilma, et dokumendi sisu edastamisel muutuks. PDF-i toetavad paljud rakendused, sealhulgas veebibrauserid, sõnumsiderakendused ja loomulikult spetsiaalsed PDF-i lugejad nagu näiteks Adobe Acrobat. (ISO, 2020)

PDF on oma iseloomu poolest mõeldud valmis failide jagamiseks ja vaatamiseks. PDF-tarkvarade pakutavad redigeerimisfunktsioonid hõlmavad peamiselt märkuste tegemist, kommenteerimist, vormide täitmist, joonistamist ja visuaalsete objektide lisamist. Tasub ka märkida, et enamik rakendusi, mis lisaks PDF-faili vaatamisele seda ka redigeerida lubavad, on tasulised. (ISO, 2020)



### 3.2. DOCX

Suurema osa oma ajaloost on Microsoft Word kasutanud salvestatud failide jaoks patenteeritud vormingut DOC. Alates 2007. aastast värskendati vaikesalvestusvormingut ning DOC'ist sai DOCX. (Microsoft, 2023)

Microsoft Office 2007 väljalaskega muudeti selle uue dokumendivormingu struktuur tavalisest kahendvormingust XML- ja binaarfailide kombinatsiooniks (mida ka lühendile lisatud X tähistab). See tähendab, et kui varasemalt koguti kõik dokumendiga seonduvad andmed ühte faili, siis nüüd on andmed jaotatud erinevate XML failide vahel, mis loovad kokkupakituna DOCX faili. (Library Of Congress, 2019)

Tänu XML vormingust tingitud väikesele mahule on DOCX faile lihtsam salvestada ja toetada. Arvuti suudab korraga töödelda rohkem DOCX-faile kui vanad DOC-failid ja DOCX-i kasutamine võimaldab arvuti töötlemissüsteemil kiiremini töötada. (Library Of Congress, 2019)

DOCX failid salvestavad andmed eraldi failide ja kaustadena pakitud ZIP konteinerisse. XML vorming on loodud dokumendi sisu avatumaks ja juurdepääsetavamaks muutmiseks. Näiteks dokumendi tekst salvestatakse lihttekstifailidesse ja dokumendikujutised salvestatakse üksikute pildifailidena, nagu .JPG või .PNG .DOCX failis. Kaasatud failid võivad sisaldada ka lehe vorminguteavet, autorluse andmeid ja dokumendi märkusi. (Yakovenko, 2016)

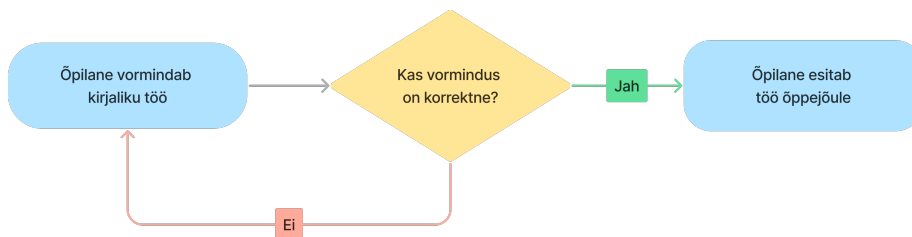
DOCX vormingus määratletud struktuurid pakuvad tekstitöötlusdokumendile laiendatud XML-sõnavara. Laiendatud elemendid ja atribuudid võimaldavad vormingul näidata lisateavet dokumendi kohta või näidata dokumendi osade sisu ja vormingut. Kuna need elemendid ja atribuudid on mõeldud laiendustena, on nende eesmärk ja kasutus erinev. (Microsoft, 2023)

### 3.3. Järeldus

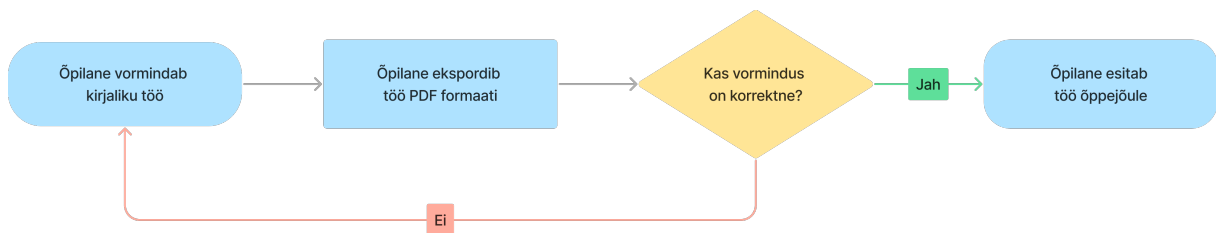
Kuna kirjalike tööde automaatkontroll oleks eelkõige tööriistaks üliõpilastele, kellest enamik (88%) kasutab kirjalike tööde kirjutamisel ja vormindamisel Microsoft Office Word'i, siis on mõistlik rakendust arendades lähtuda samast tarkvarast ning selle failiformaadist (DOCX).

Microsoft Office Word lubab üliõpilastel kohe selt vajalikud vormindusmuudatused sisse viia ning vähendab sellega kirjalike tööde esitamise elutsükli keerukust (joonis 5).

Alljärgnevatel joonistel on näha, et juhul kui automaatkontroll revideeriks PDF-faili (joonis 6), siis tekitaks see üliõpilasele kirjalike tööde vormindamisel ja edastamisel lisasamme. Nimelt ei saa õpilased vajalikke vormindusvigu PDF tarkvarades parandada, mistõttu peaksid nad pärast DOCX faili PDF'iks konverteerimist taaskord Microsoft Office Wordi muudatusi sisse viima suunduma:



Joonis 5. DOCX automaatkontroll



Joonis 6. PDF automaatkontroll

#### 4. VORMINDUSNÕUDED

Järgnevas peatükis on kirjeldatud Haapsalu kolledži kirjalike tööde juhendit ning kõrvutatud seda dokumendi parameetritega.

Vormindusnõuded on selle uurimuse raames laias laastus kaheks jagatud: muutuvad ja muutumatud. Muutuvate nõuete all on mõeldud kirjeid, mis sõltuvad töö laadist, autorist, ajast või juhendajast. See tähendab, et sõltuvalt tööst võib üks või teine nõue omada erinevat sisu.

Näide nõudest: „*Tiitelleht kajastab töö juhendaja andmeid (paremal servas lühendatult teaduslik kraad lühendatult ning ees-ja perekonnanimi). Kui tööl on mitu õppejõudu, juhendajat ja/või konsultanti, siis kirjutatakse nende andmed üksteise alla. Kui juhendaja/konsultant on väljaspoolt kõrgkooli, võib lisada nime ette ära ka töökoha ja ameti.*“ (Haapsalu kolledž, 2022)

Antud nõude puhul on väga palju lubatud variatsioone erineva sisuga alates tõsiasjast, et igal tööl ei ole juhendajat ja lõpetades olukorraga, kus tööl on üks juhendaja kõrgkoolist ning üks väljaspool kõrgkooli.

Muutumatud nõuded on näiteks eeldused lehe formaadile, kirjatüübile, reavahele jms. See tähendab, et sõltumata kirjalikust tööst on need nõuded alati ühed ja samad.

Näide nõudest: „*Tekst kirjutatakse lehe ühele küljele reasammuga 1,5.*“ (Haapsalu kolledž, 2022)

Antud nõue on alati sama igal kirjalikul tööl olenemata sellest, kas tegemist on referaadi või seminaritööga.

Selle jaotuse järgi on mõistlik automaatselt kontrollida vaid muutumatuid nõudeid, sest muutuvate nõuete puhul saaks kontrollida vaid tõsiasja, et vastavale kohale failis on tehtud mingi kanne (võib olla ka suvaline tekst, sest rakendus ei ole võimeline ilma variatsioone teadmata vahet tegema). Selle kontrollimine on aga autori hinnangul visuaalsele kontrollile samaväärne, mistõttu jäetakse taolised nõuded hetkel välja.

#### 4.1. Nõuete kõrvutamine dokumendi parameetritega

Järgnevas peatükis tutvustatakse nõuete ja parameetrite kõrvutamise protsessi. Kirjeldatakse, kuidas DOCX failiformaadi lahtipakkimisel vajalike failideni jõuda, milliseid rakendusi kasutada ning kuidas seoseid luua.

##### 4.1.1. Faili lahtipakkimine

DOCX failid salvestavad andmed eraldi failide ja kaustadena pakitud ZIP konteinerisse. DOCX sisule ligipääsemiseks on vaja fail esmalt lahti pakkida. Lahtipakkimiseks sobivad ZIP ekstraktorid.

DOCX struktuuri määrab ECMA-376 standard Open Packaging Convention (*OPC*) (Microsoft, 2018). Failiformaadi lahtipakkimisel tekivad järgmised kaustad:

1) \_rels, 2) customXml, 3) docProps, 4) word.

Peamised vormindust kirjeldavad failid asuvad kaustas „word“ ning kannavad nime document.xml ja styles.xml.

##### 4.1.2. Faili lugemine

XML faile saab lugeda ka tavalise tekstiredaktori või brauseriga, kuid antud töös on failide mugavamaks käsitlemiseks kasutatud laialtlevinud rakendust Visual Studio Code. Visual Studio Code (VS Code) on lihtne lähtekoodi redaktor, mis on vabavarana saadaval nii Windows, macOS kui ka Linux operatsioonisüsteemiga arvutitele (Microsoft, 2023).

Järgnevalt on välja toodud document.xml failis leiduvad dokumendi formaati kirjeldavad read (koodinäide 1).

```
<w:headerReference w:type="default" r:id="rId9" />  
<w:pgSz w:w="11907" w:h="16839" w:code="9" />
```

Koodinäide 1. VS Code's avatud document.xml'i dokumendi formaati kirjeldavad read

### 4.1.3. Faili kõrvutamine

Nõuete kõrvutamiseks DOCX failiformaadi struktuuri parameetritega kasutati Office Open XML failivormingute standardit ECMA-376 (joonis 8). See dokumentatsioon määratleb Office Open XML sõnavara ja dokumentide esituse ning pakkimise. (ECMA, 2016)

#### 17.6.13 pgSz (Page Size)

This element specifies the properties (size and orientation) for all pages in the current section.

[Example: Consider a section that must be printed on A4 paper. The WordprocessingML for this paper size is as follows:

```
<w:pgSz w:w="11907" w:h="16839" />
```

Joonis 8. ECMA näide dokumendi formaati kirjeldavast parameetrist (ECMA, 2016)

## 4.2. Olulised tähelepanekud

Järgnevas peatükis on välja toodud tähelepanekuid protsessi ja võrdlustabelite paremaks mõistmiseks. Kirjeldatakse ka Microsoft Office mõõtühikute loogikat ja seletatakse lahti DOCX stiilihierarhia.

### 4.2.1. Mõõtühikud

Parameetreid võrreldes selgus, et Microsoft Word kasutab oma DOCX failiformaatides mõõtühikuna peamiselt dxa'd. Dxa on Microsoft Wordi standardne mõõtühik ja see vastab ühele kahekümnendikule punktist ( $1 \text{ dxa} = \frac{1}{20} \text{ pt}$ ) (ECMA, 2016).

Loodud tabelis nõuete ning DOCX struktuuri kõrvutamiseks on märgata ka arvutusi, mille kohaselt on dxa konverteeritud ülemaailmselt kasutatavatesse tollidesse:

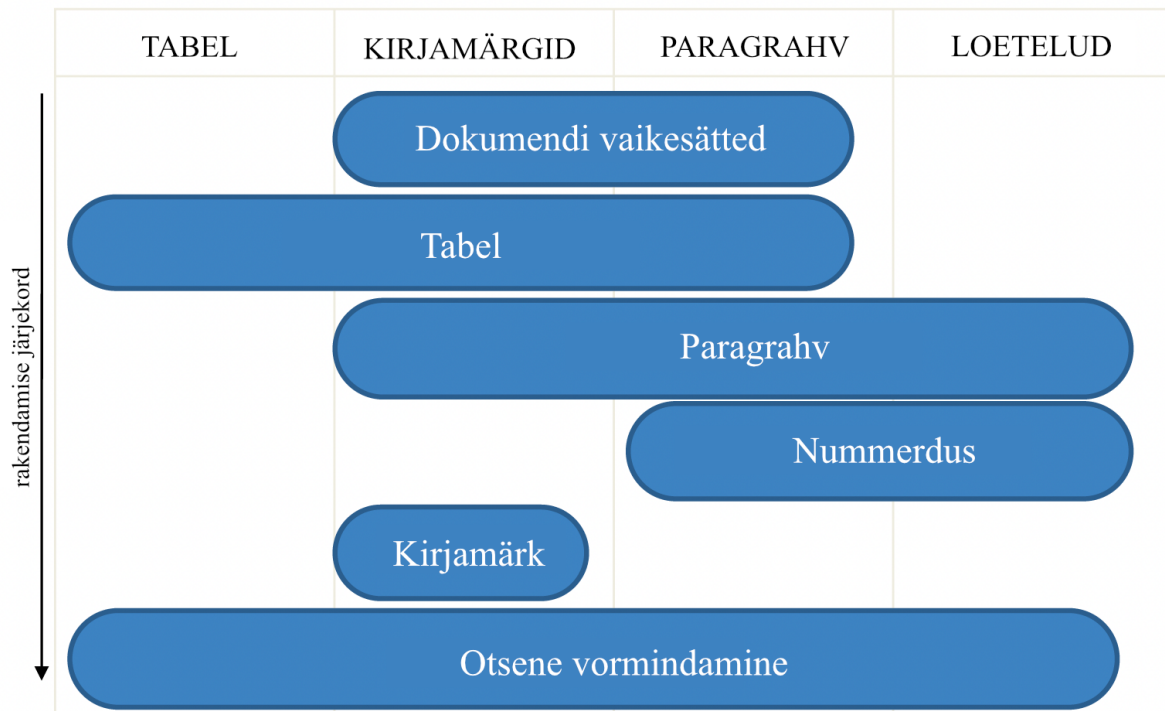
$$1 \text{ dxa} = \frac{1}{20} \text{ pt} = \frac{1}{1440} \text{ in.}$$

### 4.2.2. Stiilihierarhia

Wordis leiduva stiilivalikuga (joonis 9) võib ühes failis mitut stiilitüüpi samale sisule rakendada. Seetõttu on loodud kindel reeglistik atribuutide rakendamise järjekorraks (joonis 9).

Seejuures on oluline teada, et ühe stiilitüübiga määratud vormindamisomadusi saab muuta, eemaldada, või üle kirjutada. (ECMA, 2016)

Järgmine joonis illustreerib stiilitüüpide rakendamise järjekorda ja atribuute, mida need mõjutavad.



Joonis 9. DOCX failiformaadi stiilihierarhia rakendamise järjekord

Protsessi etappe kirjeldatakse järgmiselt (ECMA, 2016):

1. Dokumendile kehtetatakse vaikesätted, mis rakenduvad kõikidele dokumendi kirjamärkidele ja paragrahvidele;
2. Igale dokumendi tabelile rakendatakse tabelistiili atribuudid, järgides vormingutingimusi ka tabelile määratud välistuste kohta;
3. Igale paragrahvile rakendatakse nummerdatud üksuse ja paragrahvi atribuudid, mis on vormindatud nummerdamisstiiliga;
4. Igale paragrahvile rakendatakse tema stiilis määratletud paragrahvi- ja käitamisomadusi;

5. Igale osale kohaldatakse erinevaid käitumisatribuute, millele on omistatud spetsiifiline märgistiil;
6. Rakendatakse otsevorming (paragrahvi või osade omadused, mitte stiilid). Kui otsene vormindamine sisaldab nummerdamist, rakendatakse numeratsiooni ja sellega seotud paragrahvi atribuute.

Teadmised .DOCX faili stiilihierarhiast on võimaliku automaatkontrolli rakenduse arenduse aspektist olulised. Seda eelkõige seetõttu, et eelnevalt käsitletud atribuutide rakendamise järjekord annab juhised vajalike stiilide kontrollimiseks ning nende leidmiseks erinevates stiilitüüpides.

## 5. PARAMEETRID

Haapsalu kolledži kirjalike tööde juhendi nõuete kõrvutamiseks koodis esinevate parameetritega loodi tabelid. Kolm erinevat tabelit (Lisa 1., Lisa 2., Lisa 3.) jagavad eri teemade nõuded kategooriatesse. Andmete tabeli kujul esitamist peeti mõistlikuks, sest sellisel viisil on edasise võimaliku arenduse jaoks kogu vajalik info hõlpsalt kokku koondatud.

Järgneval leheküljel on toodud näiteks osa tabelist (Lisa 1.), mida uurimuse käigus täideti (Tabel 1). Tabeli esimeses veerus on nõue Haapsalu kolledži kirjalike tööde juhendist. Tabeli teises veerus parameeter DOCX koodist, mis antud nõuet kirjeldab ja kolmandas veerus parameetrit kirjeldav kommentaar kui seda vajalikuks peeti. Tabeli viimases veerus on märgitud DOCX failiformaadi vastavasisuline algdokument (XML) ning sellekohane viide ECMA-376 dokumentatsiooni.

Nagu eelnevalt töös kirjeldatud, valiti uurimistöö skoopi muutumatud nõuded. Parameetreid võrreldes jõuti aga selgusele, et mõne nõude spetsiifilisust ei ole ECMA dokumentatsioonis parameetritena kajastatud. Selliste nõuete puhul on autor lisanud kommentaari lahtrisse oma visiooni võimalikust lahendusest otsesid ECMA suunitlusi kasutamata.

Lisaks defineerimata parameetritele saadi uurimuse käigus aru, et mitmete nõuete kontrollimine on visuaalselt vähe aeganõudev. Selliste nõuete puhul ei pruugi kontrolli automatiseerimise mõju olla tasakaalus arendusele kuluva ressursiga. Tabelites on sellekohased read värvitud kollaseks.



Tabel 1. Lehekülje kujunduse nõuded ja neile vastavad parameetrid

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Töö vormistamisel kasutatakse valget vertikaalasendis lehte formaadiga A4 (210 x 297 mm) portrait asetusega.	<w:pgSz w:w="11907" w:code="9" w:h="16839"/>	Kui w:w="11907"(width) ja w:h="16839"(height) siis on tegemist portrait asetusega. 11907*0.01764≈210mm	document.xml ECMA-376 17.6.13
Tekst kirjutatakse lehe ühele küljele reasammuga 1,5.	<w:pPr> <w:spacing w:lineRule="auto" w:line="360" w:after="240" w:before="240"/>	"w:line" atribuut tähistab paragrahvi sees olevate tekstiridade vertikaalset vahet Kui w:lineRule atribuut = "auto", siis interpreeritakse reavahe "w:line" atribuudi väärtus 240ndik "w:line" väärtusest ehk 360:240=1.5	styles.xml, ECMA-376 17.3.1.33
Kirjatüüp: Times New Roman	<w:rPr> <w:rFonts w:cs="Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:ascii="Times New Roman"/>	Koodis on esindatud kolm levinuimat tekstiandmete märgikodeeringuvormingut. Sellest tulenevalt on kirjatüüp neile kõigile eraldi defineeritud.	styles.xml, ECMA-376 17.3.2.26
Tähesuurus 12 pt.	<w:rPr> <w:sz w:val="24"/> <w:szCs w:val="24"/>	Elementide sz ja szCs "w:val" atribuudi väärtus määratakse <i>half-pointides</i> (pool-punktides).	styles.xml, ECMA-376 17.3.2.38 17.3.2.39

## 6. VÕIMALIKUD LAHENDUSED KONTROLLI AUTOMATISEERIMISEKS

Järgnevas peatükis kirjeldatakse ühte võimalikku automaatkontrolli rakenduse töövoogu ning esitatakse lihtne visioon etappide kontrollimise algoritmist. Peatükis arutletakse ka, kas automaatkontrolli realiseerimiseks võiks kasutada tööjaama rakendust, MS Word laiendit või veebirakendust.

### 6.1. Töövoog

Võimalik etappide järjestus rakenduse kontrolli teostamiseks:

1. Etapp – dokumendi sisselugemine;
2. Etapp – dokumendi lahtipakkimine;
3. Etapp – dokumendi parameetrite kõrvutamine;
4. Etapp – tagasiside andmine.

Esmalt peaks rakendus sisse lugema faili (DOCX), millele vorminduskontrolli soovitakse teostada. Nagu uurimistööst selgus, siis asuvad vormindusreegleid sisaldavad XML failid DOCX konteineris. Seetõttu tuleb peale dokumendi sisselugemist see lahti pakkida, et vajalike algfailideni (syles.xml, document.xml) jõuda.

Olles dokumendi sisselugenud ning lahtipakkinud peaks rakendus tingimustele vastavaid vormindusnõudeid sisseloetud dokumendiga võrdlema. Selleks peaks rakendusel olema eelnevalt defineeritud parameetritega fail (nt. rules.xml), mis on loodud selle töö viienda peatüki tabeli baasil ning sisaldab Haapsalu kolledži kirjalike tööde vorminduse stiilikirjeldust. Kuna vormindusnõuded võivad muutuda, siis peaks see fail olema hõlpsasti muudetav.

Nõudeid võrreldes peaks rakendus lähtuma stiilihierarhiast ning dokumendi viitest (kust algdokumendist mingi nõue tuleb, kas document.xml või styles.xml). Ka need tingimused peavad olema rakendusele eelnevalt defineeritud.

Viimasena peaks rakendus olema võimeline väljastama tagasiside, mis annab kasutajale märku tema vorminduseksimustest. Tagasiside andmisel peaks rakendus lähtuma nii sisseloetud failist

kui ka eelnevalt defineeritud reeglitest. Näiteks ütleks rakendus: „Dokumendi formaat on X, aga peab olema Y“.

## 6.2. Võimalik realiseerimine

Taolise rakenduse realiseerimiseks on mitmeid variante. Automaatkontroll võiks olla töölaua rakendus, MS Word laiendus või veebirakendus.

Töölaua rakenduse puhul peaks arvestama erinevate operatsioonisüsteemidega ning tõsiasjaga, et automaatkontrolli poleks võimalik ilma kasutaja nõusolekuta uuendada. See võib viia olukorrani, kus nõuded muutuvad ning kasutajate arvutites asuv rakendus ei ole relevantne.

MS Wordi laiendusena peaks automaatkontrolli testima kõikide kasutatavate Word versioonide peal ning veenduma, et plugin nende kõigiga ühildub. Ka laienduse puhul on tarvilik allalaadimine, mis igas arvutis (nt. raamatukogu) ei ole lubatud.

Eelnevat arvestades oleks autori hinnangul automaatkontrolli mõistlik realiseerida veebirakendusena. Veebirakenduse kasuks räägivad järgmised argumendid:

- Rakendus oleks lihtsasti ligipääsetav (kasutajad ei peaks automaatkontrolli teostamiseks enda tööjaama lisaprogrammi/laiendust paigaldama, rakendus oleks samuti kasutatav kõikidest arvutitest);
- Rakendus oleks keskselt uuendatav (kasutajad saavad olla kindlad, et nad kontrollivad enda tööd vastu kehtivaid nõudeid, sest uuenduste korral jõustuksid muudatused kõikidele kasutajatele ilma, et nad ise peaksid värskendusi tegema);
- Rakenduse kasutamine ja selle testimine oleks kasutajasõbralik (pidevate täienduste korral ei peaks kasutaja rakendust uuendama või uuesti alla laadima).

## 7. TULEMUSTE ANALÜÜS JA JÄRELDUSED

Olles Haapsalu kolledži kirjalike tööde vormindusnõudeid analüüsinud tõden, et taolise kontrollrakenduse puhul mängib rolli ka mõistlikkuse printsiip. Kui proovida järjestada nõuded nende tähtsusest ja seda tehes arvestada, kui kaua ühe või teise nõude täidetavuse kontrollimiseks manuaalselt aega võib kuluda, siis selgub kiirelt, et kõik nõuded pole võrdselt tähtsad.

Vormindusnõuete ja DOCX struktuuri kõrvutamisel tuginedes leian, et Haapsalu kolledži vormindusnõuetes on palju nõudeid, mida rakenduse abil võimalik või mõistlik kontrollida ei ole. Lisaks mängib selliste nõuete puhul rolli ka visuaalne aspekt, mille kohaselt on mõningaid vormindusvigu teistest oluliselt kergem tuvastada. Sellisteks on näiteks jaluse/päise olemasolu, tiitellehe kanded ja viitamine.

Küll aga on kirjalikes töödes palju selliseid nõudeid, mis töö puhul lehelt lehele üle kanduvad ning taas esinevad. Läbiva ühtse joone loomisel omavad taolised vormindusnõuded esmatähtsat rolli, sest moodustavad suurema osa kõikidest nõuetest. Sellisteks nõueteks on näiteks lehe formaat, reasamm, joondus, kirjasuurus. Pean nende kontrollimist oluliseks, sest kirjaliku töö vorminduse kontrollimisel võtab iga lehe reasammu ja kirjasuuruse ülevaatamine oluliselt rohkem aega kui näiteks ühelt tiitellehelt silmadega üle käimine.

## KOKKUVÕTE

Diplomitöö eesmärgiks oli uurida kirjalike tööde vorminduse automaatkontrolli võimalikkust. Eesmärgi täitmiseks formuleeriti uurimisküsimused ning püstitati neile vastuste saamiseks uurimistegevused.

Uurimise käigus selgus, et sarnast dokumendi automaatkontrolli lahendust hetkel olemas ei ole, mistõttu oli vaja uurida, kuidas sellist automatiseeritud kontrolli üldse teostada saaks. Uuriti üliõpilaste küsitlusest selgunud kahte enimkasutatavat failiformaati: PDF ja DOCX ning langetati analüüsijärgselt valik DOCX kasuks, millega edaspidine töö jätkus.

Kuna oluline oli kinnitada ka teema aktuaalsus ning saada kinnitust tõstatatud hüpoteesile, siis koguti võimaliku rakendusega seonduvaid ideid ka õppejõududelt. Nagu küsitlusestki selgus, siis saaksid rakendusest kasu nii üliõpilased kui ka õppejõud. Seda peamiselt ajalise faktori näol.

DOCX kõrvutamiseks nõuetega pakiti selle faili konteiner lahti ja võrreldi koodis esinevaid parameetreid ECMA dokumentatsiooniga. Leiud märgiti tabelisse ja neist tehti kokkuvõtte.

Selgus, et olemas on eri sorti nõudeid: muutuvad ja muutumatud, mis pani piirangu rakenduse poolt kontrollitavale. Lisaks selgus, et kirjalike tööde vormindusjuhendis loetletud nõuete esinemise sagedus kirjalikes töödes on erinev, mistõttu ei pruugi olla kõikide nõuete kontrollimine mõistlik.

Sellest tulenevalt võib kokkuvõtteks öelda, et kirjalike tööde automaatkontroll on võimalik, kuid mitte kõiki Haapsalu kolledži kirjalike tööde vormindusjuhendis loetletud nõudeid kattev.

## SUMMARY

The Potential for Automated Formatting Checks on Written Assignments at Haapsalu College

The thesis aimed to investigate the prospect of automating document formatting checks on written assignments. To achieve this goal, the research questions were developed, and activities were designed to obtain answers to these questions.

Research revealed that currently there is no solution in the market for automated document formatting verification. Therefore, it was necessary to study how such process automation could be accomplished. A survey among students identified the two most used file formats for written assignments, DOCX and PDF. The DOCX format was identified as dominant; therefore, this format was selected for further analysis.

Since it was essential to confirm the topic's relevance and the validity of the hypothesis made, ideas about the potential application of such automation tool were also collected from the lecturers. The survey revealed that both, students, and lecturers, would benefit from the tool, primarily in terms of saving time.

The document formatting requirements of the DOCX were analysed by comparing the parameters available in the code of the sample file against the ECMA documentation. The findings were recorded in a tabulated format and then summarised.

It was found that there are several types of formatting requirements: variable and fixed. This variability created a limitation of what the application could check. Additionally, it was identified that the requirement to follow the formatting guide for written works varied, which means that it may not be sensible to automate checking of all requirements.

Therefore, it can be concluded that automating document formatting checks on written assignments is possible. However, if applied, the automation would only cover partial Haapsalu College document formatting requirements.

## ALLIKAD

Adobe. (s.a.). PNG files. *Adobe*. [2023, mai, 04].

<https://www.adobe.com/creativecloud/file-types/image/raster/png-file.html>

Davis, A. (2023.). What Is a JPG File: Understanding the Format. *PDF Candy*. [2023, mai, 04]

<https://pdfcandy.com/blog/what-is-jpg.html>

Dick, D. (s.a.). What is OOXML? *Office Open XML*. [2023, jaanuar 5]

<http://officeopenxml.com/index.php>

ECMA. (2016). ECMA-376 Office Open XML file formats Part I. *Ecma International*. [2023, märts 20].

<https://www.ecma-international.org/publications-and-standards/standards/ecma-376/>

George, A. (2021). What Are Plugins, and How Do They Work? *Lifewire*. [2023, mai, 04]

<https://www.lifewire.com/what-are-plugins-4582189>

Haapsalu kolledž. (2022). Üliõpilaste kirjalike tööde vormistamise juhend. *Tallinna Ülikool Haapsalu kolledž*. [2022, november 10].

<https://www.tlu.ee/sites/default/files/Haapsalu%20kolled%C5%BE/Dokumendid/HK-vormindusno%C5%BD-2022%2016-08-2022.pdf>

ISO. (2020). ISO 32000-2:2020(en) Document management — Portable document format — Part 2: PDF 2.0. *Online Browsing Platform (OBP)*. [2023, märts 20].

<https://www.iso.org/obp/ui/#iso:std:iso:32000:-2:ed-2:v1:en>

Library Of Congress. (2019). Sustainability of Digital Formats. *Planning for Library of Congress Collections*. [2023, veebruar 23].

<https://www.loc.gov/preservation/digital/formats/fdd/fdd000509.shtml#:~:text=DOC%20is%20a%20subtype%20of,the%202007%20Microsoft%20Office%20system.>

Microsoft. (2023). Visual Studio Code'i laienduse kasutamine. *Microsoft Learn*. [2023, märts 21].

<https://learn.microsoft.com/et-ee/power-apps/maker/portals/vs-code-extension>

Microsoft. (2023). [MS-DOCX]: Word Extensions to the Office Open XML (.docx) File Format. *Microsoft Learn*. [2023, jaanuar 5].

<https://interoperability.blob.core.windows.net/files/MS-DOCX/%5bMS-DOCX%5d.pdf>

Microsoft. (2018). Open Packaging Conventions Fundamentals. *Microsoft Learn*. [2023, jaanuar 5].

Õunapuu, L. (2014). Kvalitatiivne ja kvantitatiivne uurimisviis sotsiaalteadustes.



## LISA 1. ÜLIÕPILASTE KÜSIMUSTIK

Tere

Mina olen Hanna. Õpin Haapsalu Kolledži rakendusinformaatika eriala kolmandal kursusel ning olen kirjutamas oma diplomitööd.

Minu eesmärgiks on uurida kirjalike tööde vormistuse automaatkontrolli võimalikkust. Järgnev küsitlus aitab mind suunata olulistes uurimisülesannetes andes mulle ülevaate tänasest kirjalike töödega seonduvast olukorrast.

Küsimustik on anonüümne ning sellele vastamiseks kulub umbes 3 minutit.

Suur tänu, et oled nõus küsimustikule vastama!

Hanna Kalavus

Rakendusinformaatika  
Haapsalu Kolledž

\* Viitab kohustuslikule küsimusele

1. Mitmendal kursusel sa õpid? \*

- ☐ Esimesel kursusel
- ☐ Teisel kursusel
- ☐ Kolmandal kursusel

2. Mis erialal sa õpid? \*

- ☐ Rakendusinformaatika
- ☐ Käsitöö tehnoloogiad ja disain
- ☐ Tervisejuht
- ☐ Liiklusohutus

3. Mitu kirjalikku tööd oled Haapsalu Kolledži õpingute ajal esitanud?

\*

- ☐ 1-5
- ☐ 6-10
- ☐ 11-15
- ☐ Rohkem kui 15

4. Milliseid alljärgnevatest kirjalikest töödest oled oma Haapsalu Kolledži õppetöös pidanud esitama? \*

- ☐ Referaat
- ☐ Uurimistöö
- ☐ Aruanne
- ☐ Essee
- ☐ Õpimapp
- ☐ Portfoolio
- ☐ Projekt
- ☐ Muu:

5. Kui tihti nõuavad õppejõud kirjalike tööde puhul nõuetele vastavat vormistust? \*

- ☐ Alati
- ☐ Tihti
- ☐ Vahel
- ☐ Harva
- ☐ Mitte kunagi

6. Millist tekstitöötlusprogrammi sa peamiselt kirjalike tööde kirjutamisel ja \*  
vormistamisel kasutad?

- ☐ Microsoft Office Word
- ☐ OpenOffice
- ☐ LibreOffice
- ☐ Google Docs
- ☐ Apple Pages
- ☐ Muu:

7. Millises formaadis esitad peamiselt oma kirjalikke töid? \*

- ☐ .docx
- ☐ .pdf
- ☐ .odf
- ☐ Muu:

## LISA 2. ÕPPEJÕUDUDE KÜSIMUSTIK

Tere

Mina olen Hanna. Õpin Haapsalu Kolledži rakendusinformaatika eriala kolmandal kursusel ning olen kirjutamas oma diplomitööd.

Minu eesmärgiks on uurida kirjalike tööde vormistuse automaatkontrolli võimalikkust. Küsitluses uurin Haapsalu Kolledži erinevate erialade õppejõududelt, milline on tänane kirjalikke töid puudutav olukord ning millised on õppejõudude mõtted seoses võimaliku automaatkontrolliga.

Olen kõikide mõtete eest väga tänulik!

Hanna Kalavus

Rakendusinformaatika  
Haapsalu Kolledž

\* Viitab kohustuslikule küsimusele

1. Mis erialale te loenguid annate? \*

- ☐ Käsitöötehnoloogiad ja disain
- ☐ Liiklusohutus
- ☐ Rakendusinformaatika
- ☐ Tervisejuht

2. Kui oluliseks peate, et teile esitatav kirjalik töö oleks nõuetele vastavalt vormistatud? \*

- ☐ Väga oluliseks
- ☐ Oluliseks
- ☐ Väheoluliseks
- ☐ Mitteoluliseks

3. Kas peate vormistusnõuetele mittevastava töö lugemist ebamugavaks? \*

- ☐ Jah  
☐ Ei  
☐ Muu:

4. Kas peate kirjalike tööde vormistuse kontrollimist (vigade kommenteerimine, õpilasele tagasisaatmine) ajakulukaks? \*

- ☐ Jah  
☐ Ei  
☐ Muu:

5. Mis teeb kirjalike tööde lugemise/kontrollimise teie jaoks enim keeruliseks? \*

---

#### Vormistuse automaatkontrolli rakendus

Kujuta ette, et on olemas rakendus, mille abil üliõpilased enne kirjaliku töö esitamist selle vormistust saavad kontrollida. Rakendus annaks näiteks märku kui õpilase kontrollitava töö veerised ja teksti joondus ei vasta Haapsalu Kolledži kirjalike tööde nõuetele. Pange tähele, et rakendus pööraks vaid tähelepanu vormistusvigadele, kuid ei parandaks neid.

6. Millised võiksid olla selle rakenduse plussid? \*

---

7. Millised võiksid olla selle rakenduse miinused? \*

---

8. Vaid eelnevale infole tuginedes, mida sellelt rakenduselt ootaksite?

---

### LISA 3. LEHEKÜLJE KUJUNDUS

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Töö vormistamisel kasutatakse valget vertikaalasendis lehte formaadiga A4 (210 x 297 mm) portrait asetusega.	<w:pgSz w:w="11907" w:code="9" w:h="16839"/>	Kui w:w="11907"(width) ja w:h="16839"(height) siis on tegemist portrait asetusega. 11907*0.01764≈210mm	document.xml, ECMA-376 17.6.13
Tekst kirjutatakse lehe ühele küljele reasammuga 1,5.	<w:pPr> <w:spacing w:lineRule="auto" w:line="360" w:after="240" w:before="240"/>	"w:line" atribuut tähistab paragrahvi sees olevate tekstiridade vertikaalset vahet Kui w:lineRule atribuut = "auto", siis interpreeritakse reavahe "w:line" atribuudi väärtus 240ndik "w:line" väärtusest ehk 360:240=1.5.	styles.xml, ECMA-376 17.3.1.33
Kirjatüüp: Times New Roman	<w:rPr> <w:rFonts w:cs="Times New Roman" w:hAnsi="Times New Roman" w:eastAsia="Times New Roman" w:ascii="Times New Roman"/>	Koodis on esindatud kolm levinuimat tekstiandmete märgikodeeringuvormingut. Sellest tulenevalt on kirjatüüp neile kõigile eraldi defineeritud.	styles.xml, ECMA-376 17.3.2.26
Tähesuurus 12 pt.	<w:rPr> <w:sz w:val="24"/> <w:szCs w:val="24"/>	Elementide sz ja szCs "w:val" atribuudi väärtus määratakse <i>half-pointides</i> (pool-punktides).	styles.xml, ECMA-376 17.3.2.38 17.3.2.39

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Lehe servadest jäetakse vasakule ja paremale 2,5 cm vaba ruumi (margins). Ülemine ja alumine veeris (margin) on 3,25 cm või enam	<w:pgMar w:gutter="0" w:footer="709" w:header="709" w:left="1418" w:bottom="1843" w:right="1418" w:top="1843"/>	Elemendi pgMar atribuutide väärtused määratakse dxa mõõtühikutes: 1 dxa= 1/1440inch=0.00176388889cm 1418*0.00176388889=~2.5cm.	styles.xml, ECMA-376 17.6.11
Tekst kujundatakse nii, et ka lehekülje parempoolne serv jääb sirge (rööpjoondus).	<w:pPr> <w:spacing w:before="240" w:after="240" w:line="360" w:lineRule="auto"/> <w:jc w:val="both"/> </w:pPr>	Defineerib teksti kahe marginaali vahele. Atribuudi "w:jc" väärtus "both" laiendab iga teksti rea vasaku ja parema veeriseni.	styles.xml ECMA-376 17.3.1.13 17.18.44
Tekstilõikude vahele jäetakse lõiguvahe 12 pt peale lõiku.	<w:p> <w:pPr> <w:spacing w:after="240" /> </w:pPr> </w:p>	240dxa/20=12pt	styles.xml ECMA-376 17.3.1.33
Lõigu alguses ei kasutata taandrida.	<w:pPr> <w:ind w:firstLine="720"/> </w:pPr>	TAB klahviga taane: pPr elementi tekkis uus atribuut "w:ind" millel parameeter "w:firstLine" w:firstLine tähendab, et taane määratakse lõigu esimesele reale. Kontrollida, et poleks document.xml failis.	ECMA-376 17.3.1.12

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Kõik leheküljed (v.a tiitelleht, lühikokkuvõtted ja lisad) saavad päise ja jaluse. Päise ja jaluse kaugus lehekülje servast on 1,25 cm.	<pre>&lt;w:sectPr w:rsidR="00291095" w:rsidSect="00FE7250"&gt;   &lt;w:headerReference w:type="default" r:id="rId9"/&gt;   &lt;w:pgSz w:w="11906" w:h="16838" w:code="9"/&gt;   &lt;w:pgMar w:top="1843" w:right="1418" w:bottom="1843" w:left="1418" w:header="709" w:footer="709" w:gutter="0"/&gt;   &lt;w:pgNumType w:start="1"/&gt;   &lt;w:cols w:space="708"/&gt; &lt;/w:sectPr&gt;</pre>	<p>Päised ja jalused asuvad eraldi failides word/footer1.xml;footer2.xml jne. Document.xml failist leida esimesed elementide “headerReference” ja “footerReference” parameetri “r:id” väärtused.</p> <p>Päise ja jaluse kaugus lehekülje servast määratakse pgMar elemendis “w:header” ja “w:footer” parameetritega:  <math>1 \text{ dxa} = 1/1440 \text{ inch} = 0.00176388889 \text{ cm}</math>  <math>708 * 0.00176388889 \approx 1.25 \text{ cm}</math></p>	<p>ECMA-376 17.10.5 17.10.2 17.6.11</p>
Päisesse kirjutatakse vasakule poole töö pealkiri ja selle alla selgituseks, millise tööga on tegemist (referaat, essee, tooteraport, diplomitöö jne).	<pre>&lt;w:jc w:val="left"/&gt; &lt;w:t&gt;Valikpraktika&lt;/w:t&gt;</pre>	Tõenäoliselt võimalik kontrollida, et päises ja tiitellehel oleks sama tekstijupp olemas.	<p>header2.xml ECMA-376 17.3.1.13</p>



Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Päisesse paremale, töö laadi nimetusega samale reale kirjutatakse töö valmimise aasta.	<pre>&lt;w:tabs&gt;   &lt;w:tab w:val="center" w:pos="4537"/&gt;   &lt;w:tab w:val="right" w:pos="8765"/&gt; &lt;/w:tabs&gt;  &lt;w:t&gt;aruanne&lt;/w:t&gt;   &lt;w:tab/&gt;   &lt;w:tab/&gt;   &lt;w:t xml:space="preserve"&gt;2023&lt;/w:t&gt;</pre>	<p>Tab elemendis on kirjeldatud ära spetsiifilised tab stopid selle headeri jaoks. Seejärel on peale nimetust kasutatud tabi kaks korda.</p> <p>Tõenäoliselt võimalik kontrollida, et headeris oleks kontrollimise hetkega sama aasta.</p>	header2.xml ECMA-376 17.3.1.38
Jalusesse kirjutatakse vasakule poole õppeasutuse nimetus (Tallinna Ülikooli Haapsalu kolledž) ja selle alla töö autori ees- ja perekonnanimi.	<pre>&lt;w:jc w:val="left" /&gt; &lt;w:t xml: space="preserve"&gt;Tallinna Ülikooli Haapsalu Kolledž&lt;/w:t&gt;</pre>	Atribuudi w:jc parameeter w:val="left" määrab vasakpoolse joonduse.	footer1.xml footer2.xml ECMA-376 17.3.1.13
Jalusesse paremale, õppeasutuse nimetusega samale reale paigutatakse leheküljenumber.	<pre>&lt;w:instrText&gt;PAGE&lt;/w:instrText&gt;</pre>	Parameeter „PAGE“ on dünaamiline väärtus, mis uuendatakse automaatselt.	footer1.xml ECMA-376 17.16 17.16.5.44

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Lehekülgede numeratsiooni hulka loetakse kõik lehed alates tiitellehest. Lehekülje numbreid peidetakse kõikidel lehekülgedel enne sisukorda ja lisades.	<w:pgNumType w:start="2"/>	Atribuudiga w:start määratakse number, mida kuvatakse esimesel lehel, millel number ei ole peidetud.	document.xml ECMA-376 17.6.12
Päises pannakse piirjoon kirjutatu alla, jaluses kirjutatu peale.	<w:drawing> ~120 rida koodi </w:drawing>	Piirjoone kuvamine tekitab faili mahuka w:drawing elemendi.	header1.xml header2.xml ECMA-376 17.3.3.9
Kirja suurus päises ja jaluses on 10 pt.	w:rPr><w:color w:val="000000"/> <w:sz w:val="20"/><w:szCs w:val="20"/></w:rPr>	-	header2.xml footer1.xml footer2.xml ECMA-376 17.3.2.38 17.3.2.39
Reasamm päises ja jaluses on 1.	<w:spacing w:before="0" w:after="0" w:line="240" w:lineRule="auto"/>	-	header2.xml footer1.xml footer2.xml ECMA-376 17.3.1.33

#### LISA 4. PEALKIRJAD

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Kõikide peatükkide, sissejuhatuse, kokkuvõtte, allikate loetelu ja lisade pealkirjad kirjutatakse läbivalt suurtähtedega.	<code>&lt;w:style w:type="paragraph" w:styleId="Heading1"&gt;   &lt;w:name w:val="heading 1"/&gt;   &lt;w:next w:val="Normal"/&gt;   &lt;w:b/&gt;   &lt;w:caps/&gt; &lt;/w:style&gt;</code>	Parameeter w:caps määrab, et vastavat „Heading“ stiili omav tekst on kirjutatud suurtähtedega.	styles.xml ECMA-376 17.3.2.5
Alajaotuse, alapunktide pealkirjad kirjutatakse väiketähtedega (v.a suur algustäht).	-	Eraldi parameetrit suure algustähe jaoks ei ole.	-

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Kõikide pealkirjade puhul kasutatakse paksu (Bold) kirjastiili.	<pre>&lt;w:style w:type="paragraph" w:styleId="Heading2"&gt;   &lt;w:name w:val="heading 2"/&gt;   &lt;w:next w:val="Normal"/&gt;   &lt;w:link w:val="Heading2Char"/&gt;   &lt;w:uiPriority w:val="9"/&gt;   &lt;w:unhideWhenUsed/&gt;   &lt;w:qFormat/&gt;   &lt;w:rsid w:val="00F362A5"/&gt;   &lt;w:pPr&gt;     &lt;w:keepNext/&gt;     &lt;w:keepLines/&gt;     &lt;w:spacing w:before="480"/&gt;     &lt;w:outlineLvl w:val="1"/&gt;   &lt;/w:pPr&gt;   &lt;w:rPr&gt;     &lt;w:b/&gt;     &lt;w:color w:val="000000"/&gt;   &lt;/w:rPr&gt; &lt;/w:style&gt;</pre>	Parameeter w:b määrab, et vastavat „Heading#“ stiili omav tekst on kirjutatud suurtähtedega.	styles.xml ECMA-376 17.3.2.1
Peatükkide ja teiste alljaotuste pealkirjade järel punkti ei panda.	-	Eraldi parameetrit järgneva punkti kontrollimise jaoks ei ole.	-
Pealkiri paigutatakse lehe vasakusse serva.	<w:jc w:val="both"/>	Rööpjoonduse puhul alustatakse rida vasakult.	styles.xml ECMA-376 17.3.1.13

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Kõiki esimese taseme peatükke ja töö iseseisvaid struktuuriüksusi (sisukord, sissejuhatus, kokkuvõtte, allikate loetelu jne) alustatakse töös uult lehelt.	<pre>&lt;w:pStyle w:val="Heading1"/&gt; &lt;/w:pPr&gt; &lt;w:bookmarkStart w:id="21" w:name="_Toc123660728"/&gt; &lt;w:r&gt; &lt;w:lastRenderedPageBreak/&gt; &lt;w:t&gt;LISA 1. TELLIJ NÕUDED&lt;/w:t&gt;</pre>	“Heading1” stiili omav paragrahv peab sisaldama atribuuti w:lastRenderedPageBreak enne w:t(teksti) atribuuti.	document.xml ECMA-376 17.3.3.13
Peatüki ning alapeatüki pealkirja ja sellele järgneva teksti vahele jäetakse 12 p vahe.	<pre>&lt;w:pPrDefault&gt; &lt;w:pPr&gt; &lt;w:spacing w:before="240" w:after="240" w:line="360" w:lineRule="auto"/&gt; &lt;w:jc w:val="both"/&gt; &lt;/w:pPr&gt; &lt;/w:pPrDefault&gt;</pre>	Atribuut w:spacing w:after="240" määrab 12 p vahe. $240/20=12$ .	styles.xml ECMA-376 17.3.1.33

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Alapealkirja ning eelneva teksti vahele jäetakse 24 p vahe.	<w:style w:type="paragraph" w:styleId="Heading2"> <w:name w:val="heading 2"/> <w:next w:val="Normal"/> <w:pPr> <w:spacing w:before="480"/> <w:outlineLvl w:val="1"/> </w:pPr> <w:rPr> <w:b/> <w:color w:val="000000"/> </w:rPr> </w:style>	Atribuut w:spacing w:before="480" määrab 24 p vahe. $480/20=24$ .	document.xml ECMA-376 17.3.1.33

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Töö põhiosa (peatükid ja nende alljaotused) nummerdatakse, soovitavalt hierarhilise numeratsiooniga araabia numbritega Peatüki numbri järel pannakse punkt ning üks tühik	-	Eraldi parameetrit järgneva punkti ja tühiku kontrollimise jaoks ei ole.	-
Nummerdamata jäetakse sissejuhatuse, kokkuvõtte, allikate loetelu, registrite ja lisade peatükid.	-	Nummerdamata lehekülgede sektsioonidega seotud jaluse failid ei tohi sisaldada atribuuti <w:instrText>PAGE</w:instrText>.	footer2.xml ECMA-376 17.16 17.16.5.44

## LISA 5. TABELID, JOONISED, ILLUSTRATSIOONID JA FOTOD

Nõue	Vastav parameeter koodis	Kommentaar	Viide dokumentatsioonile
Tabeli kohale lehe vasakusse serva kirjutatakse üldnimetus Tabel	<w:jc w:val="both"/>	Enne w:tbl atribuuti olev paragrahv peab olema joondatud vasakule või rööpjoondatud. Lisaks peab paragrahvis olev teksti atribuut <w:t> sisaldama teksti „Tabel“.	document.xml ECMA-376 17.4.37 17.3.1.13 17.3.3.31
Tabel on lehe suhtes tsentreeritud.	<w:tblPr> <w:jc w:val="center" /> </w:tblPr>	Tabeli atribuutides on määratud tabeli asend lehe suhtes w:jc atribuudiga väärtusega w:val=center“.	document.xml ECMA-376 17.4.28
Üldnimetusele “Tabel” järgneb tabeli number. Numbri järele pannakse punkt.	-	Eraldi parameetrit järgneva numbri ja punkti kontrollimise jaoks ei ole.	-
Joonised, fotod ja illustratsioonid peavad paiknema lehekülje keskel	<w:jc w:val="center"/>	Pic:blipFill atribuuti sisaldav paragrahv peab samuti sisaldama w:jc w:val atribuute väärtusega „center“.	document.xml ECMA-376 20.4.2.5 20.1.8.14 17.3.1.13
Joonise numeratsioon ja nimetus kirjutatakse joonise alla.	<w:t xml:space="preserve">Joonis<w:t>	Peale pic:blipFill atribuuti olevas paragrahvis peab olema teksti atribuut <w:t>, mis sisaldab teksti „Joonis“.	document.xml ECMA-376 17.3.3.31
Üldnimetusele järgneb number. Numbri järele pannakse punkt.	-	Eraldi parameetrit järgneva numbri ja punkti kontrollimise jaoks ei ole.	-